Bacillus niameyensis sp. nov., a new bacterial species isolated from human gut

M. Tidjani Alou[1], J. Rathored[1], S. I. Traore[1], S. Khelaifia[1], C. Michelle[1], S. Brah[2], B. A. Diallo[3], D. Raoult[1,4] and J.-C. Lagier[1]

1) *Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, Inserm 1095, Faculté de médecine, Marseille, France,* 2) *Hopital National de Niamey,* 3) *Laboratoire de Microbiologie, Département de Biologie, Université Abdou Moumouni de Niamey, Niamey, Niger and* 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Bacillus niameyensis sp. nov. strain SIT3^T (= CSUR P1266 = DSM 29725) is the type strain of *B. niameyensis* sp. nov. This Gram-positive strain was isolated from the digestive flora of a child with kwashiorkor and is a facultative anaerobic rod and a member of the *Bacillaceae* family. This organism is hereby described alongside its complete genome sequence and annotation. The 4 286 116 bp long genome (one chromosome but no plasmid) contains 4130 protein-coding and 66 RNA genes including five rRNA genes.

Keywords: *Bacillus niameyensis*, culturomics, genome, malnutrition, taxonogenomics
Original Submission: 8 July 2015; **Revised Submission:** 13 September 2015; **Accepted:** 14 September 2015
Article published online: 28 September 2015

Corresponding author: J.-C. Lagier, Aix-Marseille Université, URMITE, UMR CNRS 7278, IRD 198, INSERM U1095, Faculté de Médecine, 27 Bd Jean Moulin, 13385 Marseille Cedex 5, France
E-mail: jclagier@yahoo.fr
M. Tidjani Alou, J. Rathored and S.I. Traore contributed equally to this article, and all should be considered first author.

Introduction

Bacillus niameyensis strain SIT3^T (= CSUR P1266 = DSM 29725) is the type strain of *B. niameyensis* sp. nov. This bacterium is a Gram-positive bacillus, spore-forming, facultative anaerobic, and motile. It was isolated from the stool of a child living in Niamey, Niger, afflicted with kwashiorkor. This isolation was part of a culturomics study of the gut microbiota of children with severe acute malnutrition aiming to characterize their microbiota. Culturomics aims to explore as exhaustively as possible a microbial ecosystem using multiple culture conditions [1,2].

Phylogenetic relationships based on the 16S ribosomal RNA gene are currently used for bacterial classification alongside phenotypic and genotypic characteristics [3–5]. However, with the development of genomic sequencing and its decreasing cost, a new concept of bacterial description has been used in our laboratory [2,6–11] combining a proteomics description with the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) profile [12] associated with a biochemical and genomic description of the new bacterial species.

The genus *Bacillus* was established in 1872 by Cohn and encompasses over 200 described species and subspecies belonging to the *Firmicutes* phylum. *Bacillus* species are strictly aerobic and facultative anaerobic rod-shaped bacteria that form heat-resisting endospores [12–15]. They are often found in the environment (water, soil, air), food, plants and human clinical samples [16]. Some *Bacillus* species, such as *Bacillus thuringiensis,* are known to be pathogenic for insects and are used as biological control agents for crops [14].

Materials and methods

Organism information: classification and features

A stool sample was collected from a 2-year-old boy living in Niamey, Niger, with kwashiorkor, a form of severe acute

http://dx.doi.org/10.1016/j.nmni.2015.09.011

malnutrition. Consent was obtained from the child's parents at the National Hospital of Niamey and the study was approved by the Institut Fédératif de Recherche 48, Faculty of Medicine, Marseille, France, under agreement 09-022. The patient did not receive antibiotics at the time of sample collection, and the fecal sample was stored at −80°C.

Strain identification

The stool sample was cultured using the culturomics concept [2]. MALDI-TOF was used for colony identification as described below. In case of a failed identification using this technique, the 16S ribosomal RNA was sequenced. Stackebrandt and Ebers [17] suggested that a similarity level lower than 98.7% defined a new species (Fig. 1) without performing DNA-DNA hybridization (DDH). This similarity value is below the 16S rRNA threshold of 98.7% set by Stackebrandt and Ebers to delineate a new species without carrying out DDH.

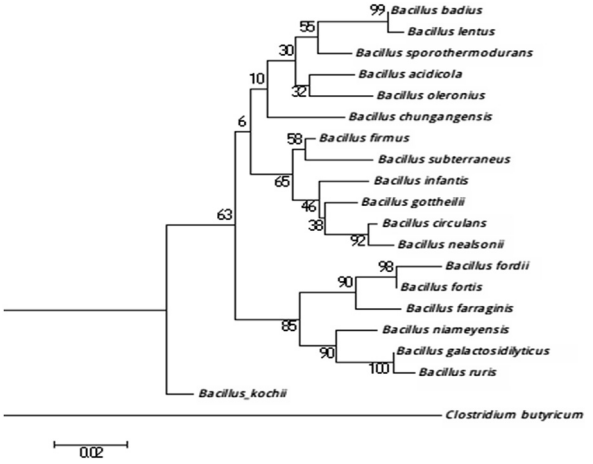

FIG. 1. Phylogenetic tree highlighting position of *Bacillus niameyensis* sp. nov. strain SIT3 (= CSUR P1266 = DSM 29725) relative to other type strains within *Bacillus* genus. Strains and their corresponding GenBank accession numbers for 16S rRNA genes are *Bacillus amyloliquefaciens* AB006920, *Bacillus bataviensis* AJ542508, *Bacillus drentensis* AJ542506, *Bacillus endophyticusv* AF295302, *Bacillus galactosidilyticus* AJ535638, *Bacillus idriensis* AY904033, *Bacillus mojavensis* AB021191, *Bacillus infantis* AY904032, *Bacillus niacini* AB021194, *Bacillus novalis* AJ542512, *Bacillus pseudomycoides* AF13121, *Bacillus pocheonensis* AB245377, *Bacillus thuringiensis* D16281, *Bacillus thermoamylovorans* L27478. Sequences were aligned using Clustal W (http://www.clustal.org/clustal2/), and phylogenetic inferences were obtained using maximum-likelihood method within MEGA 6 (http://www.megasoftware.net/mega.php). Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 1000 times to generate majority consensus tree. *Clostridium butyricum* AJ458420 was used as outgroup. Scale bar = 1% nucleotide sequence divergence.

Phenotypic characteristics

Phenotypic characteristics like Gram staining (Fig. 2), sporulation, motility, catalase and oxidase were highlighted as previously described [18]. Biochemical features of our strain were investigated using API 20NE, ZYM and 50 CH strips according to the manufacturer's instructions (bioMérieux, Marcy l'Étoile, France). Various growth temperatures (25, 30, 37, 45 and 56°C) and atmospheres were tested. Growth under anaerobic and microaerophilic conditions occurrence was tested using GENbag Anaer and GENbag miroaer systems, respectively (bioMérieux). Aerobic growth was achieved with or without 5% CO_2.

In order to perform the electronic microscopy to observe our strain (Fig. 3), detection Formvar-coated grids were deposited on a 40 µL bacterial suspension drop and incubated during 30 minutes at 37°C. The grids were incubated for 10 seconds on ammonium molybdate 1%, dried on blotting paper and then observed with a Morgani 268D transmission electron microscope (Philips/FEI, Hillsboro, OR, USA) at an operating voltage of 60 kV.

MALDI-TOF protein analysis was carried out as previously described [12,19] using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). Twelve individual colonies were deposited on a MTP 96 MALDI-TOF target plate (Bruker). The 12 spectra were imported into the MALDI BioTyper software (version 2.0, Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 6.252 bacteria, including 199 spectra from 104 validly named *Bacillus* species used as reference data in the BioTyper database. A score enabled the presumptive identification and discrimination of the tested species from those in a database: a score of >2 with a validated species enabled the identification at the species level, and a score of <1.7 did not enable any identification. No significant score was obtained for strain SIT3^T, thus

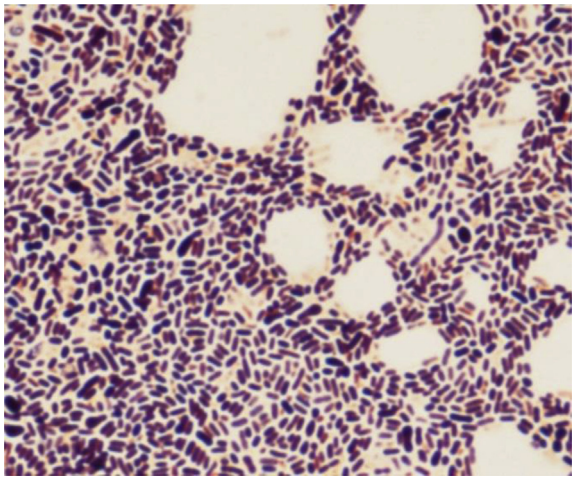

FIG. 2. Gram staining of *B. niameyensis* strain SIT3^T.

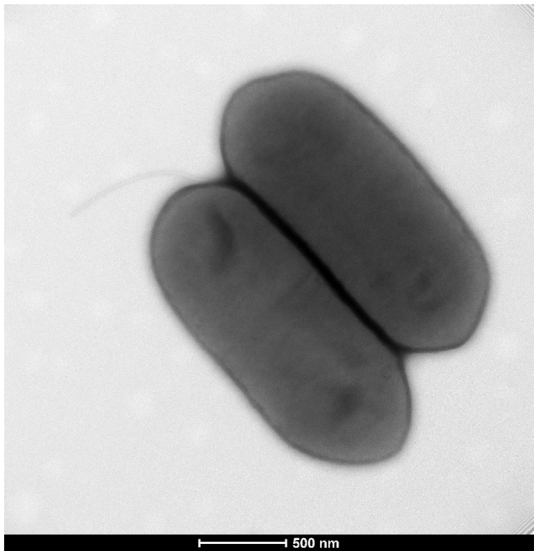

FIG. 3. Transmission electron microscopy of *B. niameyensis* strain SIT3, using Morgani 268D transmission electron microscope (Philips/FEI, Hillsboro, OR, USA) at operating voltage of 60 kV. Scale bar = 500 nm.

suggesting that our isolate was not a member of any known species. The reference spectrum for strain SIT3^T (Fig. 4) was incremented in our database and then compared to other known species of the *Bacillus* genus. The differences are shown in Fig. 5.

Growth conditions and genomic DNA preparation

B. niameyensis strain SIT3^T (= CSUR P1266 = DSM 29725) was grown on 5% sheep's blood–enriched Columbia agar (bioMérieux) at 37°C in aerobic atmosphere. Bacteria grown on three petri dishes were collected and resuspended in 4 × 100 µL of Tris-EDTA (TE) buffer. Then 200 µL of this suspension was diluted in 1 mL TE buffer for lysis treatment that included a 30-minute incubation with 2.5 µg/µL lysozyme at 37°C, followed by an overnight incubation with 20 µg/µL proteinase K at 37°C. Extracted DNA was then purified using three successive phenol–chloroform extractions and ethanol precipitations at −20°C overnight. After centrifugation, the DNA was resuspended in 160 µL TE buffer.

Genome sequencing and assembly

Genomic DNA of *Bacillus niameyensis* was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA was bar coded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with a high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 66.2 ng/µL. The mate pair library was prepared with 1 µg of genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 lab chip. The DNA fragments ranged in size from 1 to 11 kb, with an optimal size at 3.927 kb. No size selection was performed, and 505 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 597 bp on a Covaris device S2 (Covaris, Woburn, MA, USA) in microtubes. The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured at 59.2 nmol/L.

The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the

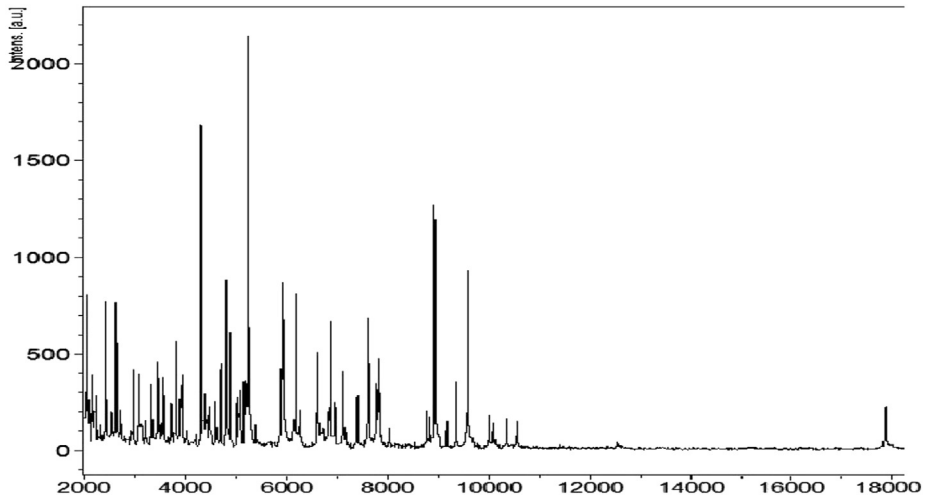

FIG. 4. Reference mass spectrum from *B. niameyensis* strain SIT3^T. Spectra from 12 individual colonies were compared and reference spectrum generated.

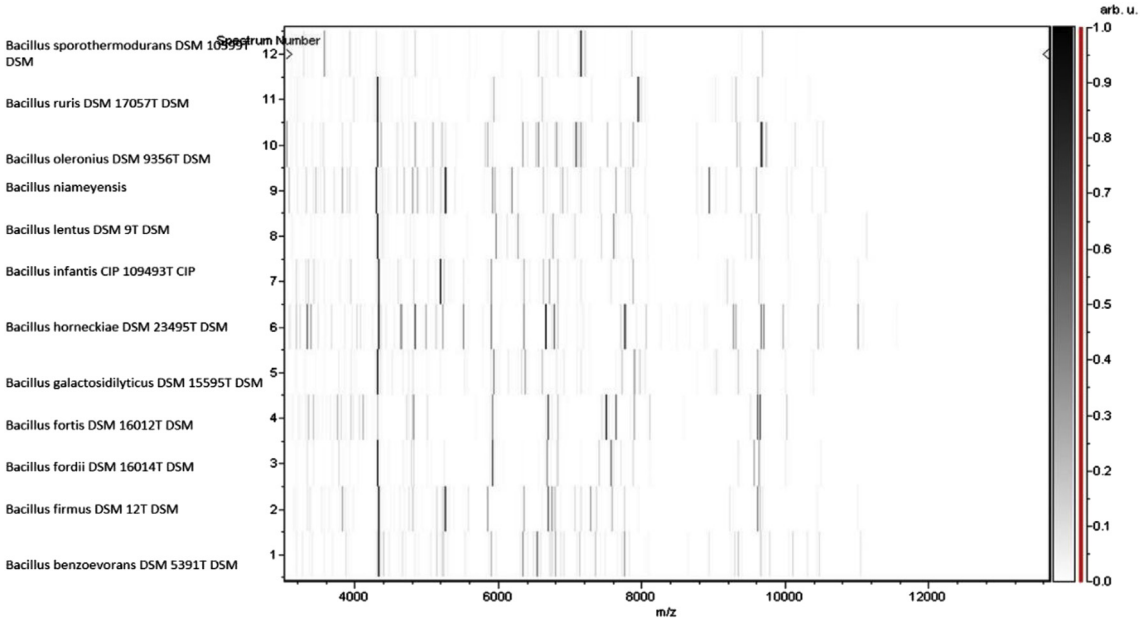

FIG. 5. Gel view comparing *B. niameyensis* (= CSUR P1266 = DSM 29725) to other species within genus *Bacillus*. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. X-axis records m/z value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity expressed by greyscale scheme code. Color bar and right y-axis indicate relation between color peak, with peak intensity in arbitrary units. Displayed species are indicated at left.

instrument along with the flow cell. An automated cluster generation and sequencing run was performed in a single 39-hour run in a 2 × 251 bp.

Genome annotation and genome analysis

Open reading frames (ORFs) were predicted using Prodigal with default parameters, but the predicted ORFs were excluded if they were spanning a sequencing gap region (containing N). The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COGs) database using BLASTP (E value $1e^{-03}$, coverage 0.7 and identity percentage 30%). If no hit was found, it was searched against the NR database using BLASTP with an E value of $1e^{-03}$, coverage 0.7 and identity percentage 30%. If the sequence length was smaller than 80 aa, we used an E value of $1e^{-05}$. The tRNAScanSE tool [20] was used to find tRNA genes, whereas ribosomal RNAs were found by using RNAmmer [21]. Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [22]. ORFans were identified if all the BLASTP analyses performed did not provide positive results (E value smaller than $1e^{-03}$ for ORFs with sequence size larger than 80 aa or E value smaller than $1e^{-05}$ for ORFs with sequence length smaller than 80 aa). Such parameter thresholds have already been used in previous studies to define ORFans.

Genomes were automatically retrieved from the 16S RNA tree using Xegen software (Phylopattern) [23]. For each selected genome, complete genome sequence, proteome genome sequence and Orfeome genome sequence were retrieved via File Transfer Protocol (FTP) from the National Center for Biotechnology Information. All proteomes were analysed with proteinOrtho [24]. Then for each couple of genomes, a similarity score was computed. This score is the mean value of nucleotide similarity between all couple of orthologues between the two genomes studied (average genomic identity of orthologous gene sequences, AGIOS) [25]. An annotation of the entire proteome was performed to define the distribution of functional classes of predicted genes according to the

TABLE 1. **Classification and general features of *Bacillus niameyensis* strain SIT3^T**

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Firmicutes* Class: *Bacilli* Order: *Bacillales* Family: *Bacillaceae* Genus: *Bacillus* Species: *Bacillus niameyensis* Type strain: SIT3
Gram stain	Positive
Cell shape	Rod
Motility	Motile
Sporulation	sporulating
Temperature range	Mesophilic
Optimum temperature	37°C

TABLE 2. Differential characteristics of *Bacillus niameyensis* strain SIT3^T, *Bacillus hackensackii, Bacillus galactosidilyticus* DSM 15595, *Bacillus fordii* DSM 16014, *Bacillus fortis* DSM 16012, *Bacillus oceanisediminis* JCM 16506, *Bacillus sporothermodurans* DSM 10599, *Bacills infantis* JCM 13438, *Bacillus horneckiae* MTCC 9535, *Bacillus lentus* ATCC 10840 [29–36]

Property	*B. niameyensis*	*B. hackensackii*	*B. galactosidilyticus*	*B. fordii*	*B. fortis*	*B. oceanisediminis*	*B. sporothermodurans*	*B. infantis*	*B. horneckiae*	*Bacillus lentus*
Cell diameter (μm)	0.7–1.0	2.0–8.0	0.7–0.9	0.6–0.8	0.6–0.8	0.6–0.8	0.7	NA	1.0–1.5	0.7–1.2
Oxygen requirement	+	+	+	+	+	+	+	+	+	+
Gram stain	+	+	+	–	–	+	+	+	+	+
Salt requirement	NA	NA	NA	NA	NA	NA	NA	NA	–	–
Motility	+	+	+	+	+	NA	+	NA	+	+
Endospore formation	+	+	+	+	+	+	+	NA	+	+
Indole	–	–	–	–	–	+	NA	NA	–	NA
Production of										
Alkaline phosphatase	–	NA	NA	NA	NA	NA	NA	NA	+	NA
Catalase	+	+	+	+	+	+	+	+	+	+
Oxidase	–	–	NA	+	+	+	+	–	–	NA
Nitrate reductase	+	–	+	NA	NA	+	–	NA	+	–
Urease	–	+	+/–	NA	NA	–	–	NA	–	+
β-Galactosidase	+	NA	+	NA	NA	+	NA	NA	–	NA
N-acetyl-glucosamine	+	NA	+	+/–	+	NA	NA	+	+	NA
Acid from:										
L-Arabinose	–	–	+/–	w	–	–	NA	NA	+	+
Ribose	–	–	+/–	w	+	–	NA	+	–	NA
Mannose	–	–	+	–	–	–	–	–	+	+
Mannitol	–	–	–	–	–	+	–	+	+	+
Sucrose	–	–	+/–	–	–	+	NA	+	–	+
D-Glucose	–	–	w	–	+	+	–	+	–	+
D-Fructose	–	–	w	–	–	–	–	+	–	NA
D-Maltose	+	–	w	–	–	–	NA	+	+	NA
D-Lactose		–	+/–	–	–	–	–	+	–	NA
Habitat	Human gut	Blood	Raw milk	Raw milk	Milking apparatus	Marine sediment	Milk	Blood	Spacecraft assembly	soil

NA, data not available; w, weak reaction.

clusters of orthologous groups of proteins (using the same method than for the genome annotation). To evaluate the genomic similarity among studied *Bacillus* strains, we determined two parameters, dDDH – digital DDH (DNA-DNA Hybridization), which exhibits a high correlation with DDH [24,25], and AGIOS [23], which was designed to be independent from DDH.

Results

Phenotypic description

Strain SIT3^T (Table 1) was first isolated in May 2014 by a 15-day preincubation in blood culture with sheep blood and cultivation on 5% sheep blood–enriched Colombia agar (bioMérieux) in aerobic atmosphere at 37°C. This strain showed a 95.1% nucleotide sequence similarity with *Bacillus lentus,* the phylogenetically closest *Bacillus* species with a valid published name (Fig. 1).

Growth occurred between 28 and 45°C on blood-enriched Columbia agar (bioMérieux), with the optimal growth being obtained at 37°C after 48 hours of incubation. Optimal growth was achieved aerobically. Weak cell growth was observed under microaerophilic and anaerobic conditions. The motility test was positive, and the cells were sporulating. Colonies were translucent and 1 to 2 mm in diameter on sheep's blood–enriched Columbia agar. Cells were Gram-positive rods (Fig. 2). Under electron microscopy, the bacteria grown on agar had a mean diameter and length of 0.87 μm and 2.6 μm, respectively (Fig. 3).

Strain SIT3^T exhibited catalase activity but was negative for oxidase.

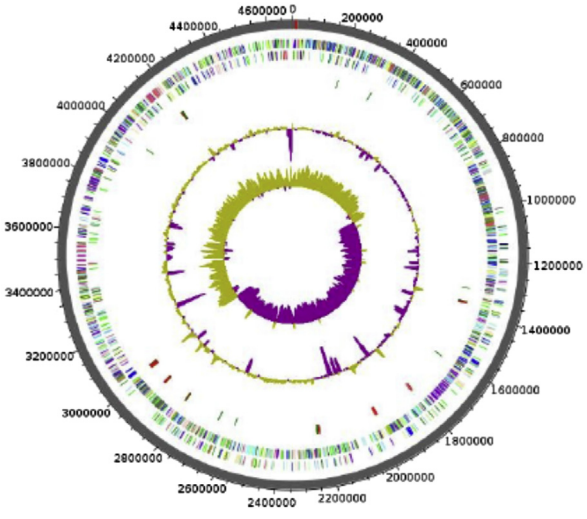

FIG. 6. Graphical circular map of chromosome. From outside to center: genes on forward strain colored by COGs categories (only gene assigned to COGs), RNA genes (tRNAs green, rRNAs red), GC content and GC skew. COGs, Clusters of Orthologous Groups database.

Using an API ZYM strip (bioMérieux), positive reactions were observed for esterase (C4), esterase lipase (C8), leucine arylamidase, α-chymotrypsin, naphtol-AS-BI-phosphohydrolase, β-galactosidase and N-acetyl-β-glucosaminidase. Negative reactions were observed for acid phosphatase, lipase (C14), valine arylamidase, cysteine arylamidase, trypsin, alkaline phosphatase, α-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-fucosidase and α-mannosidase.

Using an API 20 NE strip, positive reactions were obtained for nitrate reduction, β-glucosidase and β-galactosidase. All other reactions were negative.

Using an API 50 CH strip (bioMérieux), positive reactions were recorded for methyl-αD-glucopyranoside, N-acetylglucosamine, esculin ferric citrate, D-maltose, D-trehalose and starch. Negative reactions were recorded for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-βD-xylopranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-αD-mannopyranoside, amygdaline, arbutin, salicin, D-cellobiose, D-lactose, D-melibiose, D-saccharose, inulin, D-raffinose, D-melezitose, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate.

Cells are susceptible to imipenem, metronidazole, doxycycline, rifampicin, vancomycin, amoxicillin, ceftriaxone and gentamicin (500 µg) and are resistant to trimethoprim/sulfamethoxazole, erythromycin, ciprofloxacin and gentamicin (15 µg).

TABLE 3. Nucleotide content and gene count levels of the genome

Attribute	Genome (total)	
	Value	% of total[a]
Size (bp)	4 286 116	100
G + C content (bp)	1 603 007	37.4
Coding region (bp)	3 665 154	85.51
Total genes	4196	100
RNA genes	66	1.57
Protein-coding genes	4130	98.42
Genes with function prediction	3058	72.88
CRISPRs	7	0.16
Genes assigned to COGs	2694	64.20
Genes with peptide signals	294	7.0
Genes with transmembrane helices	1180	28.12
Genes associated to PKS or NRPS	10	0.23
Genes associated to mobilome	2198	51.80
Genes associated to toxin/antitoxin	111	2.61
Genes associated to resistance genes	0	0
Genes with paralogues (E value $1e^{-10}$)	1317	31.03
Genes larger than 5000 nt	2	0.05

COGs, Clusters of Orthologous Groups database; CRISPR, clustered regularly interspaced short palindromic repeat; NRPS, nonribosomal peptide synthase; PKS, polyketide synthase.
[a]Total is based on either size of genome (bp) or total number of protein coding genes in annotated genome.

TABLE 4. Number of genes associated with the 25 general COGs functional categories

Code	Value	% of total[a]	Description
J	180	4.36	Translation
A	0	0.00	RNA processing and modification
K	297	7.19	Transcription
L	185	4.48	Replication, recombination and repair
B	1	0.02	Chromatin structure and dynamics
D	36	0.87	Cell cycle control, mitosis and meiosis
Y	0	0.00	Nuclear structure
V	111	2.69	Defense mechanisms
T	186	4.50	Signal transduction mechanisms
M	173	4.19	Cell wall/membrane biogenesis
N	64	1.55	Cell motility
Z	0	0.00	Cytoskeleton
W	0	0.00	Extracellular structures
U	44	1.07	Intracellular trafficking and secretion
O	99	2.40	Posttranslational modification, protein turnover, chaperones
C	156	3.78	Energy production and conversion
G	351	8.50	Carbohydrate transport and metabolism
E	279	6.76	Amino acid transport and metabolism
F	86	2.08	Nucleotide transport and metabolism
H	97	2.35	Coenzyme transport and metabolism
I	85	2.06	Lipid transport and metabolism
P	218	5.28	Inorganic ion transport and metabolism
Q	73	1.77	Secondary metabolites biosynthesis, transport and catabolism
R	520	12.59	General function prediction only
S	326	7.89	Function unknown
—	374	8.91	Not in COGs

COGs, Clusters of Orthologous Groups database.
[a]Total is based on either size of genome (bp) or total number of protein coding genes in annotated genome.

The differences exhibited by the comparison with other representatives of the genus *Bacillus* are detailed in Table 2.

Genome properties

The genome of *B. niameyensis* strain SIT3^T is 4 286 116 bp long with a 37.40% G + C content (Fig. 6, Table 3) and 46 generated contigs. Of the 4196 predicted genes, 4130 were protein-coding genes and 66 were RNAs. Of these 66 rRNA genes, one is a 16S rRNA gene, three are 5S RNA genes, one is a 23S rRNA and 61 predicted tRNA genes were identified in the genome. A total of 3068 genes (73.11%) were assigned a putative function. A total of 230 genes were identified as ORFans (5.42%). Using ARG-ANNOT [26], no resistance genes were found. Nevertheless, ten genes associated to polyketide synthase (PKS) or nonribosomal peptide synthase (NRPS) were discovered in the analysis of the genome. The remaining genes were annotated as hypothetical proteins. The properties and

TABLE 5. Closely related species with *Bacillus niameyensis*

Name of species	Similarity (%)	Accession No.
Bacillus niameyensis	100	LK985389
Bacillus oceanisediminis	94.88	KP120947.1
Bacillus nealsonii	94.45	KC433938.1
Bacillus firmus	93.50	KF478238.1
Bacillus infantis	94.60	KP696715.1
Bacillus fordii	95.56	NR_025786.1

TABLE 6. Number of orthologous proteins shared between genomes (upper right)[a]

	Bacillus niameyensis	*Bacillus badius*	*Bacillus firmus*	*Bacillus fordii*	*Bacillus infantis*	*Bacillus nealsonii*	*Bacillus oceanisediminis*	*Clostridium butyricum*
Bacillus niameyensis	4,130	1,511	1,769	1,648	1,800	1,658	1,846	832
Bacillus badius	61.24	4,486	1,940	1,730	1,886	1,601	2,040	838
Bacillus firmus	61.8	62.60	4,142	1,956	2,634	2,062	3,039	971
Bacillus fordii	65.06	61.82	62.01	4,688	1,842	1,608	2,069	824
Bacillus infantis	59.99	62.25	67.18	61.3	4,142	2,074	2,728	946
Bacillus nealsonii	61.44	60.15	62.84	59.95	61.08	4,789	2,149	957
Bacillus oceanisediminis	61.82	62.67	88.49	62.03	67.17	63.04	5,578	1,012
Clostridium butyricum	54.21	51.46	52.78	52.44	50.8	54.49	52.8	4,152

[a]Average percentage similarity of nucleotides corresponding to orthologous proteins shared between genomes (lower left) and numbers of proteins per genome.

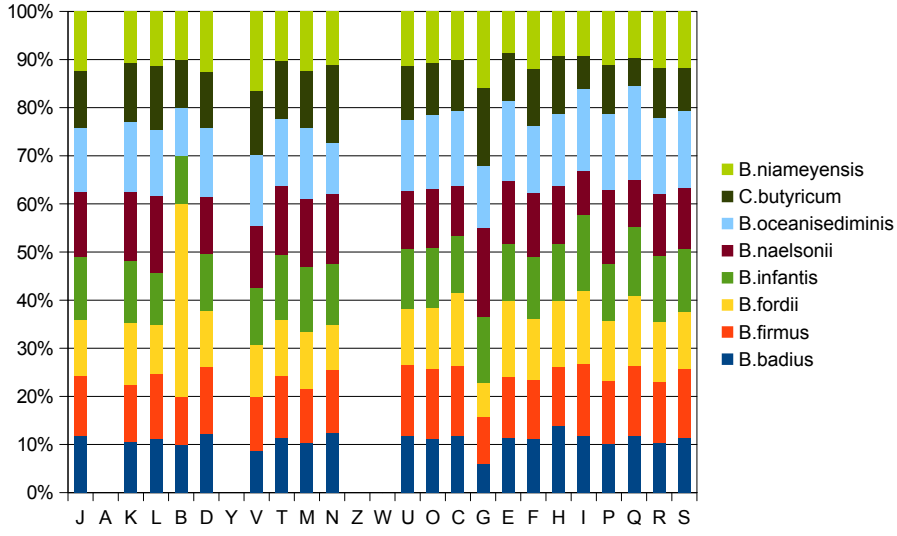

FIG. 7. Distribution of functional classes of predicted genes according to clusters of orthologous groups of proteins.

statistics of the genome are summarized in Table 3. The distribution of genes into COGs functional categories is presented in Table 4.

Genome comparison

Bacillus oceanideminis, Bacillus nealsonii, Bacillus firmus, Bacillus infantis and *Bacillus fordii* are closely related species to *B. niameyensis* with available genomes (Table 5) and were thus chosen for this comparative analysis. The draft genome sequence of *Bacillus niameyensis* is smaller than those of *Bacillus oceanisediminis, Bacillus nealsonii, Bacillus firmus, Bacillus infantis* and *Bacillus fordii* (5.76, 4.98, 4.97, 4.88, 4.62 and 4.51 MB, respectively) but larger than those of *Bacillus badius* (4.04 MB). It should be noted that the size of our genome is an estimation

TABLE 7. Pairwise comparison of *Bacillus niameyensis* with eight other species using GGDC, formula 2 (DDH estimates based on identities/HSP length)[a] [27,28]

	Bacillus niameyensis	*Bacillus badius*	*Bacillus firmus*	*Bacillus fordii*	*Bacillus infantis*	*Bacillus nealsonii*	*Bacillus oceanisediminis*	*Clostridium butyricum*
Bacillus niameyensis	100% ± 00	2.53% ± 0.18	2.53% ± 0.20	2.54% ± 0.22	2.54% ± 0.13	2.54% ± 0.20	2.53% ± 0.22	2.53% ± 0.17
Bacillus badius		100% ± 00	2.54% ± 0.19	2.53% ± 0.17	2.55% ± 0.18	2.53% ± 0.17	2.54% ± 0.21	2.53% ± 0.19
Bacillus firmus			100% ± 00	2.53% ± 0.19	2.60% ± 0.22	2.55% ± 0.20	3.04% ± 0.11	2.53% ± 0.18
Bacillus fordii				100% ± 00	2.53% ± 0.18	2.53% ± 0.17	2.53% ± 0.19	2.52% ± 0.18
Bacillus infantis					100% ± 00	2.54% ± 0.15	2.60% ± 0.21	2.53% ± 0.18
Bacillus nealsonii						100% ± 00	2.56% ± 0.16	2.53% ± 0.18
Bacillus. oceanisediminis							100% ± 00	2.53% ± 0.18
Clostridium butyricum								100% ± 00

DDH, DNA-DNA hybridization; GGDC, genome-to-genome distance; HSP, high-scoring pair.
[a]Confidence intervals indicate inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size). These results are in accordance with the 16S rRNA (Fig. 1) and phylogenetic analyses as well as GGDC results.

because 46 contigs were obtained in the sequencing of the genome.

The G + C content of *Bacillus niameyensis* is smaller than those of *Bacillus infantis, Bacillus badius, Bacillus firmus, Bacillus fordii* and *Bacillus oceanisediminis* (46.0, 43.90, 41.40, 41.20 and 40.8%, respectively) but larger than those of *Bacillus nealsonii* and *Clostridium butyricum* (35.1 and 28.8%).

The gene content of *Bacillus niameyensis* is smaller than those of *Bacillus oceanisediminis Bacillus nealsonii, Bacillus infantis, Bacillus badius, Bacillus fordii, Bacillus firmus* and *Clostridium butyricum* (5722, 4864, 4837, 4486, 4229, 4229 and 4231, respectively). However the distribution of genes into COGs categories was similar in all compared genomes (Table 6, Fig. 7). In addition, *Bacillus niameyensis* shared 4130, 4486, 4142, 4688, 4142, 4789, 5578 and 4152 orthologous genes with *Bacillus badius, Bacillus firmus, Bacillus fordii, Bacillus infantis, Bacillus nealsonii, Bacillus oceanisediminis* and *Clostridium butyricum* (Table 7). Among species with standing in nomenclature, AGIOS values ranged from 61.24 between *Bacillus niameyensis* and *Bacillus badius* to 63.04% between *Bacillus oceanisediminis* and *Bacillus nealsonii.*

Conclusion

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of *Bacillus niameyensis* sp. nov. that contains the strain SIT3^T. This bacterial strain was isolated from the fecal flora of a 2-year-old boy from Niamey, Niger, with kwashiorkor.

Taxonomic and nomenclatural proposals

Description of *Bacillus niameyensis* strain SIT3^T sp. nov. Cells are Gram-positive, sporulating, rod-shaped bacilli with a diameter of 0.1 µm. Colonies are translucent and 1 to 2 mm in diameter on 5% sheep's blood–enriched Columbia agar (bioMérieux). Cells are catalase positive and oxidase negative. Positive reactions were observed for esterase (C4), esterase lipase (C8), leucine arylamidase, α-chymotrypsin, naphtol-AS-BI-phosphohydrolase, β-galactosidase and N-acetyl-β-glucosaminidase. Negative reactions were observed for acid phosphatase, lipase (C14), valine arylamidase, cysteine arylamidase, trypsin, alkaline phosphatase, α-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-fucosidase and α-mannosidase.

Cells are susceptible to imipenem, metronidazole, doxycycline, rifampicin, vancomycin, amoxicillin, ceftriaxone and gentamycin (500 µg) and are resistant to trimethoprim/sulfamethoxazole, erythromycin, ciprofloxacin and gentamicin (15 µg).

The G + C content of the genome is 37.40%. The 16S rRNA gene sequence and whole-genome shotgun sequence of *B. niameyensis* strain SIT3^T are deposited in GenBank under accession numbers LK955389 and CTDY01000000, respectively. The type strain SIT3^T (= CSUR P1266 = DSM 29725) was isolated from the stool of a child living in Niamey, Niger, with kwashiorkor.

Conflict of interest

None declared.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process. This study was funded by the Méditerranée Infection Foundation. We thank Karolina Griffiths for English-language review.

References

[1] Dubourg G, Lagier JC, Armougom F, et al. The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. Eur J Clin Microbiol Infect 2013;32:637–45.
[2] Lagier JC, Armougom F, Million M, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[3] Viale AM, Arakaki AK, Soncini FC, et al. Evolutionnary relationships among eubacterial groups as inferred from GroEL (Chaperonin) sequences comparison. Int J Syst Bacteriol 1944;44:527–33.
[4] Woese CR, Kandler O, Wheelis ML, et al. Towards a natural system of organisms: proposals for the domains *Archaea, Bacteria*, and *Eucarya*. Proc Natl Acad Sci U S A 1990;87:4576–9.
[5] Wolf M, Muller T, Dandekar T, et al. Phylogeny of firmicutes with special reference to *Mycoplasma (Mollicutes)* as inferred from phosphoglycerate kinase amino acid sequence data. Inst J Syst Evol Microbiol 2004;54(pt 3):871–5.
[6] Hugon P, Mishra AK, Lagier JC, et al. Non-contiguous finished genome sequence and description of *Brevibacillus massiliensis* sp. nov. Stand Genomic Sci 2013;8:1–14.
[7] Kokcha S, Ramasamy D, Lagier JC, et al. Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. Stand Genomic Sci 2012;7:233–45.
[8] Lagier JC, El Karkouri K, Nguyen TT, et al. Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. Stand Genomic Sci 2012;6:116–25.
[9] Lagier JC, Armougom F, Mishra AK, et al. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. Stand Genomic Sci 2012;6:315–24.
[10] Lagier JC, Elkarkouri K, Rivet R, et al. Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. Stand Genomic Sci 2013;7:343–56.

[11] Lagier JC, El Karkouri K, Mishra AK, et al. Non contiguous-finished genome sequence and description of *Enterobacter massiliensis* sp. nov. Stand Genomic Sci 2013;7:399–412.
[12] Seng P, Drancourt M, Gouriet F, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543–51.
[13] Fritze D. Taxonomy of the genus *Bacillus* and related genera: the aerobic endospore-forming bacteria. Physiopathology 2004;94: 1245–8.
[14] Gordon RE, Haynes WC, Pang CH, et al. The genus *Bacillus*. US Dep Agric Handh 1973;427.
[15] Priest FG, Goodfellow M, Todd C, et al. A numerical classification of the genus *Bacillus*. J Gen Microbiol 1988;134:1847–82.
[16] Goto K, Omura T, Hara Y, Sadaie Y, et al. Application of the 16S rDNA sequence as an index for rapid identification of species in the genus *Bacillus*. J Gen Appl Microbiol 2000;46:1–8.
[17] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[18] Lagier JC, Edouard S, Pagnier I, Mediannikov O, Drancourt M, Raoult D. Current and past strategies for bacterial culture in clinical microbiology. Clin Microbiol Rev 2015;28:208–36.
[19] Seng P, Abat C, Rolain JM, et al. Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.
[20] Bendtsen JD, Nielsen H, von Heijne G, et al. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[21] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997;25:955–64.
[22] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338: 1027–36.
[23] Gouret P, Thompson JD, Pontarotti P. PhyloPattern: regular expressions to identify complex patterns in phylogenetic trees. BMC Bioinform 2009;10:298.
[24] Lechner M, Findeib S, Steiner L, et al. Proteinortho: detection of (co-) orthologs in large-scale analysis. BMC Bioinform 2011;12:124.
[25] Ramasamy D, Mishra AK, Lagier JC, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[26] Gupka SK, Padmanabhan BR, Diene SM, et al. ARG-ANNOT, a new bioinformatics tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58:212–20.
[27] Auch AF, Von Jan M, Klenk HP, et al. DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–34.
[28] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 2013;14:60.
[29] Hong T, Heibler N, Tang Y, et al. *"Bacillus hackensackii"* sp. nov., a novel carbone dioxide sensitive bacterium isolated from blood culture. Diagn Microbiol Infect Dis 2003;45:143–7.
[30] Heyndrickx M, Logan NA, Lebbe L, et al. *Bacillus galactosidilyticus* sp. nov., an alkali-tolerant beta-galactosidase producer. Int J Syst Evol Microbiol 2004;54:617–21.
[31] Scheldeman P, Rodriguez-Diaz M, Goris J, et al. Heyndrickx M. *Bacillus farraginis* sp. nov., *Bacillus fortis* sp. nov., *Bacillus fordii* sp. nov., isolated from dairy farms. Int J Syst Evol Microbiol 2004;54:1355–64.
[32] Zhang J, Wang J, Fang C, et al. *Bacillus oceanisediminis* sp. nov., isolated from marine sediment. Int J Syst Evol Microbiol 2010;60:2924–9.
[33] Pettersson B, Lembke F, Hammer P, et al. *Bacillus sporothermodurans*, a new species producing highly heat-resistant endospores. Int J Syst Evol Microbiol 1996;46:759–63.
[34] Ko KS, Oh WS, Lee MY, Lee, et al. *Bacillus infantis* sp. nov. and *Bacillus idriensis* sp. nov., isolated from a patient with a neonatal sepsis. Int J Syst Evol Microbiol 2006;56:2541–4.
[35] Vaishampayan P, Probst A, Krishnamurthi S, et al. *Bacillus horneckiae* sp. nov., isolated from a spacecraft-assembly clean room. Int J Syst Evol Microbiol 2010;60:1031–7.
[36] Gordon RE, Hyde JL, Moore Jr JA. *Bacillus firmus–Bacillus lentus*: a series or one species? Int J Syst Bacteriol 1977;27:256–62.